

SEQUENCE LISTING

(1) INFORMATION FOR SEQ ID NO: 1:

(i)SEQUENCE CHARACTERISTICS:
 (A)LENGTH: 6 amino acids
 (B)TYPE: amino acid
 (D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(v)FRAGMENT TYPE: internal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Asn Asp Gln Val Leu Phe
1 5

(2) INFORMATION FOR SEQ ID NO: 2:

(i)SEQUENCE CHARACTERISTICS:
 (A)LENGTH: 6 amino acids
 (B)TYPE: amino acid
 (D)TOPOLOGY: linear

(ii)MOLECULE TYPE: internal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Phe Glu Asp Met Thr Asp
1 5

(3) INFORMATION FOR SEQ ID NO: 3:

(i)SEQUENCE CHARACTERISTICS:
 (A)LENGTH: 7 amino acids
 (B)TYPE: amino acid
 (D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(v)FRAGMENT TYPE: internal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Phe Lys Leu Ile Leu Lys Lys
1 5

(4) INFORMATION FOR SEQ ID NO: 4:

(i)SEQUENCE CHARACTERISTICS:
 (A)LENGTH: 5 amino acids
 (B)TYPE: amino acid
 (D)TOPOLOGY: linear

(ii)MOLECULE TYPE: internal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Tyr Lys Asp Ser
1 5

(5) INFORMATION FOR SEQ ID NO: 5:

(i)SEQUENCE CHARACTERISTICS:
(A)LENGTH: 5 amino acids
(B)TYPE: amino acid
(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: internal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Ser Thr Leu Ser Cys
1 5

(6) INFORMATION FOR SEQ ID NO: 6:

(i)SEQUENCE CHARACTERISTICS:
(A)LENGTH: 157 amino acids
(B)TYPE: amino acid
(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn	
1				5				10						15		
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp	
			20					25					30			
Met	Thr	Asp	Ser	Asp	Cys	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile	
		35					40					45				
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile	
	50					55					60					
Ser	Val	Lys	Cys	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Cys	Glu	Asn	Lys	Ile	
65					70					75				80		
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys	
				85					90					95		
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys	
			100					105					110			
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Cys	Glu	
		115					120					125				
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu	
	130					135					140					
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp				
145				150						155						

(7) INFORMATION FOR SEQ ID NO: 7:

(i)SEQUENCE CHARACTERISTICS:
(A)LENGTH: 157 amino acids

(B)TYPE: amino acid
(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Asn	Phe	Gly	Arg	Leu	His	Cys	Thr	Thr	Ala	Val	Ile	Arg	Asn	Ile	Asn	
1				5					10					15		
Asp	Gln	Val	Leu	Phe	Val	Asp	Lys	Arg	Gln	Pro	Val	Phe	Glu	Asp	Met	
			20					25					30			
Thr	Asp	Ile	Asp	Gln	Ser	Ala	Ser	Glu	Pro	Gln	Thr	Arg	Leu	Ile	Ile	
		35					40					45				
Tyr	Met	Tyr	Lys	Asp	Ser	Glu	Val	Arg	Gly	Leu	Ala	Val	Thr	Leu	Ser	
	50					55					60					
Val	Lys	Asp	Ser	Lys	Met	Ser	Thr	Leu	Ser	Cys	Lys	Asn	Lys	Ile	Ile	
65					70					75					80	
Ser	Phe	Glu	Glu	Met	Asp	Pro	Pro	Glu	Asn	Ile	Asp	Asp	Ile	Gln	Ser	
				85					90					95		
Asp	Leu	Ile	Phe	Phe	Gln	Lys	Arg	Val	Pro	Gly	His	Asn	Lys	Met	Glu	
			100					105					110			
Phe	Glu	Ser	Ser	Leu	Tyr	Glu	Gly	His	Phe	Leu	Ala	Cys	Gln	Lys	Glu	
		115					120					125				
Asp	Asp	Ala	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Lys	Asp	Glu	Asn	Gly	Asp	
	130					135					140					
Lys	Ser	Val	Met	Phe	Thr	Leu	Thr	Asn	Leu	His	Gln	Ser				
145					150					155						

(8)INFORMATION FOR SEQ ID NO: 8:

(i)SEQUENCE CHARACTERISTICS:
(A)LENGTH: 471 base pairs
(B)TYPE: nucleic acid
(C)STRANDEDNESS: double
(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: cDNA

(vi)ORIGINAL SOURCE:
(A)ORGANISM: human
(G)CELL TYPE: liver

(ix)FEATURE:
(A)NAME/KEY: mat peptide
(B)LOCATION: 1..471
(C)IDENTIFICATION METHOD: E

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TAC	TTT	GGC	AAG	CTT	GAA	TCT	AAA	TTA	TCA	GTC	ATA	AGA	AAT	TTG	AAT	48
Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn	
1			5					10					15			
GAC	CAA	GTT	CTC	TTC	ATT	GAC	CAA	GGA	AAT	CGG	CCT	CTA	TTT	GAA	GAT	96
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp	
		20					25					30				
ATG	ACT	GAT	TCT	GAC	TGT	AGA	GAT	AAT	GCA	CCC	CGG	ACC	ATA	TTT	ATT	144

Met	Thr	Asp	Ser	Asp	Cys	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile		
		35					40					45					
ATA	AGT	ATG	TAT	AAA	GAT	AGC	CAG	CCT	AGA	GGT	ATG	GCT	GTA	ACT	ATC	192	
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile		
		50					55				60						
TCT	GTG	AAG	TGT	GAG	AAA	ATT	TCA	ACT	CTC	TCC	TGT	GAG	AAC	AAA	ATT	240	
Ser	Val	Lys	Cys	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Cys	Glu	Asn	Lys	Ile		
		65			70						75				80		
ATT	TCC	TTT	AAG	GAA	ATG	AAT	CCT	CCT	GAT	AAC	ATC	AAG	GAT	ACA	AAA	288	
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys		
				85					90					95			
AGT	GAC	ATC	ATA	TTC	TTT	CAG	AGA	AGT	GTC	CCA	GGA	CAT	GAT	AAT	AAG	336	
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys		
			100					105					110				
ATG	CAA	TTT	GAA	TCT	TCA	TCA	TAC	GAA	GGA	TAC	TTT	CTA	GCT	TGT	GAA	384	
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Cys	Glu		
		115					120					125					
AAA	GAG	AGA	GAC	CTT	TTT	AAA	CTC	ATT	TTG	AAA	AAA	GAG	GAT	GAA	TTG	432	
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu		
		130				135					140						
GGG	GAT	AGA	TCT	ATA	ATG	TTC	ACT	GTT	CAA	AAC	GAA	GAC				471	
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp					
				145		150				155							

(9) INFORMATION FOR SEQ ID NO: 9:

(i)SEQUENCE CHARACTERISTICS:

(A)LENGTH: 11 amino acids

(B)TYPE: amino acid

(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(v)FRAGMENT TYPE: N-terminal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met	Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser
1				5					10	

(10) INFORMATION FOR SEQ ID NO: 10:

(i)SEQUENCE CHARACTERISTICS:

(A)LENGTH: 10 amino acids

(B)TYPE: amino acid

(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(v)FRAGMENT TYPE: C-terminal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp
1				5				10	

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 13 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(v) FRAGMENT TYPE: N-terminal fragment

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg
1 5 10

(i)SEQUENCE CHARACTERISTICS:
 (A)LENGTH: 14 amino acids
 (B)TYPE: amino acid
 (D)TOPOLOGY: linear

```
(v)FRAGMENT TYPE: internal fragment
```

Thr Ile Phe Ile Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg
1 5 10

(i)SEQUENCE CHARACTERISTICS:
 (A)LENGTH: 17 amino acids
 (B)TYPE: amino acid
 (D)TOPOLOGY: linear

```
(v)FRAGMENT TYPE: internal fragment
```

Ile Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys
1 5 10 15

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 471 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii)MOLECULE TYPE: cDNA

(ix)FEATURE:

(A)NAME/KEY: mat peptide

(B)LOCATION: 1..471

(C)IDENTIFICATION METHOD: S

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TAC	TTT	GGC	AAG	CTT	GAA	TCT	AAA	TTA	TCA	GTC	ATA	AGA	AAT	TTG	AAT	48
Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn	
1				5				10						15		
GAC	CAA	GTT	CTC	TTC	ATT	GAC	CAA	GGA	AAT	CGG	CCT	CTA	TTT	GAA	GAT	96
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp	
			20					25						30		
ATG	ACT	GAT	TCT	GAC	TCT	AGA	GAT	AAT	GCA	CCC	CGG	ACC	ATA	TTT	ATT	144
Met	Thr	Asp	Ser	Asp	Ser	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile	
			35					40						45		
ATA	AGT	ATG	TAT	AAA	GAT	AGC	CAG	CCT	AGA	GGT	ATG	GCT	GTA	ACT	ATC	192
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile	
			50					55						60		
TCT	GTG	AAG	TCT	GAG	AAA	ATT	TCA	ACT	CTC	TCC	GCT	GAG	AAC	AAA	ATT	240
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Ala	Glu	Asn	Lys	Ile	
65				70										80		
ATT	TCC	TTT	AAG	GAA	ATG	AAT	CCT	CCT	GAT	AAC	ATC	AAG	GAT	ACA	AAA	288
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys	
				85										95		
AGT	GAC	ATC	ATA	TTC	TTT	CAG	AGA	AGT	GTC	CCA	GGA	CAT	GAT	AAT	AAG	336
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys	
			100											110		
ATG	CAA	TTT	GAA	TCT	TCA	TCA	TAC	GAA	GGA	TAC	TTT	CTA	GCT	TGT	GAA	384
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Cys	Glu	
			115					120						125		
AAA	GAG	AGA	GAC	CTT	TTT	AAA	CTC	ATT	TTG	AAA	AAA	GAG	GAT	GAA	TTG	432
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu	
			130					135						140		
GGG	GAT	AGA	TCT	ATA	ATG	TTC	ACT	GTT	CAA	AAC	GAA	GAC				471
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp				
145				150										155		

(15) INFORMATION FOR SEQ ID NO: 15:

(i)SEQUENCE CHARACTERISTICS:

(A)LENGTH: 10 amino acids

(B)TYPE: amino acid

(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(v)FRAGMENT TYPE: N-terminal fragment

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser
1				5				10	

(16) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 471 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
 (A) NAME/KEY: mat peptide
 (B) LOCATION: 1..471
 (C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TAC	TTT	GGC	AAG	CTT	GAA	TCT	AAA	TTA	TCA	GTC	ATA	AGA	AAT	TTG	AAT	48
Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn	
1				5					10					15		
GAC	CAA	GTT	CTC	TTC	ATT	GAC	CAA	GGA	AAT	CGG	CCT	CTA	TTT	GAA	GAT	96
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp	
			20					25					30			
ATG	ACT	GAT	TCT	GAC	TCT	AGA	GAT	AAT	GCA	CCC	CGG	ACC	ATA	TTT	ATT	144
Met	Thr	Asp	Ser	Asp	Ser	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile	
			35				40					45				
ATA	AGT	ATG	TAT	AAA	GAT	AGC	CAG	CCT	AGA	GGT	ATG	GCT	GTA	ACT	ATC	192
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile	
			50				55					60				
TCT	GTG	AAG	TCT	GAG	AAA	ATT	TCA	ACT	CTC	TCC	GCT	GAG	AAC	AAA	ATT	240
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Ala	Glu	Asn	Lys	Ile	
					70					75				80		
ATT	TCC	TTT	AAG	GAA	ATG	AAT	CCT	CCT	GAT	AAC	ATC	AAG	GAT	ACA	AAA	288
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys	
					85				90					95		
AGT	GAC	ATC	ATA	TTC	TTT	CAG	AGA	AGT	GTC	CCA	GGA	CAT	GAT	AAT	AAG	336
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys	
			100					105					110			
ATG	CAA	TTT	GAA	TCT	TCA	TCA	TAC	GAA	GGA	TAC	TTT	CTA	GCT	TCT	GAA	384
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Ser	Glu	
			115				120					125				
AAA	GAG	AGA	GAC	CTT	TTT	AAA	CTC	ATT	TTG	AAA	AAA	GAG	GAT	GAA	TTG	432
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu	
			130				135					140				
GGG	GAT	AGA	TCT	ATA	ATG	TTC	ACT	GTT	CAA	AAC	GAA	GAC				471
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp				
					145		150				155					

(17) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11464 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii)MOLECULE TYPE: genomic DNA

(vi)ORIGINAL SOURCE:

(A)ORGANISM: human

(G)CELL TYPE: placenta

(ix)FEATURE:

(A)NAME/KEY: 5' UTR

(B)LOCATION: 1..3

(C)IDENTIFICATION METHOD: E

(A)NAME/KEY: leader peptide

(B)LOCATION: 4..82

(C)IDENTIFICATION METHOD: S

(A)NAME/KEY: intron

(B)LOCATION: 83..1453

(C)IDENTIFICATION METHOD: E

(A)NAME/KEY: leader peptide

(B)LOCATION: 1454..1465

(C)IDENTIFICATION METHOD: S

(A)NAME/KEY: intron

(B)LOCATION: 1466..4848

(C)IDENTIFICATION METHOD: E

(A)NAME/KEY: leader peptide

(B)LOCATION: 4849..4865

(C)IDENTIFICATION METHOD: S

(A)NAME/KEY: mat peptide

(B)LOCATION: 4866..4983

(C)IDENTIFICATION METHOD: S

(A)NAME/KEY: intron

(B)LOCATION: 4984..6317

(C)IDENTIFICATION METHOD: E

(A)NAME/KEY: mat peptide

(B)LOCATION: 6318..6451

(C)IDENTIFICATION METHOD: S

(A)NAME/KEY: intron

(B)LOCATION: 6452..11224

(C)IDENTIFICATION METHOD: E

(A)NAME/KEY: mat peptide

(B)LOCATION: 11225..11443

(C)IDENTIFICATION METHOD: S

(A)NAME/KEY: 3' UTR

(B)LOCATION: 11444..11464

(C)IDENTIFICATION METHOD: E

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AAG ATG GCT GCT GAA CCA GTA GAA GAC AAT TGC ATC AAC TTT GTG GCA	48
Met Ala Ala Glu Pro Val Glu Asp Asn Cys Ile Asn Phe Val Ala	
-35 -30 -25	
ATG AAA TTT ATT GAC AAT ACG CTT TAC TTT ATA G GTAAGG CTAATGCCAT	98
Met Lys Phe Ile Asp Asn Thr Leu Tyr Phe Ile Ala	
-20 -15 -10	
AGAACAAATA CCAGGTTTCAG ATAAATCTAT TCAATTAGAA AAGATGTTGT GAGGTGAACT	158
ATTAAGTGAC TCTTTGTGTC ACCAAATTTC ACTGTAATAT TAATGGCTCT TAAAAAATA	218
GTGGACCTCT AGAAATTAAC CACAACATGT CCAAGGTCTC AGCACCTTGT CACACCACGT	278
GTCTTGGCAC TTTAATCAGC AGTAGCTCAC TCTCCAGTTG GCAGTAAGTG CACATCATGA	338

AAATCCCAGT	TTTCATGGGA	AAATCCCAGT	TTTCATTGGA	TTTCCATGGG	AAAAATCCCA	398
GTACAAAAC	GGGTGCATTC	AGGAAATACA	ATTTCCCAAA	GCAAATTGGC	AAATTATGTA	458
AGAGATTCTC	TAAATTTAGA	GTTCCGTGAA	TTACACCATT	TTATGTAAAT	ATGTTTGACA	518
AGTAAAAATT	GATTCTTTTT	TTTTTTTTCT	GTTGCCCAGG	CTGGAGTGCA	GTGGCACAAT	578
CTCTGCTCAC	TGCAACCTCC	ACCTCCTGGG	TTCAAGCAAT	TCTCCTGCCT	CAGCCTTCTG	638
AGTAGCTGGG	ACTACAGGTG	CATCCCCGCC	TGCCTGGCTA	ATTTTGGGT	ATTTTACTA	698
GAGACAGGGT	TTTGGCATGT	TGTCCAGGCT	GGTCTTGGAC	TCCTGATCTC	AGATGATCCT	758
CCTGGCTCGG	GCTCCCAAAG	TGCTGGGATT	ACAGGCATGA	ACCACCACAC	ATGGCCTAAA	818
AATTGATTCT	TATGATTAAT	CTCCTGTGAA	CAATTTGGCT	TCATTTGAAA	GTTTGCCTTC	878
ATTTGAAACC	TTCATTTAAA	AGCCTGAGCA	ACAAAGTGAG	ACCCCATCTC	TACAAAAAAC	938
TGCAAAATAT	CCTGTGGACA	CCTCCTACCT	TCTGTGGAGG	CTGAAGCAGG	AGGATCACTT	998
GAGCCTAGGA	ATTTGAGCCT	GCAGTGAGCT	ATGATCCCAC	CCCTACACTC	CAGCCTGCAT	1058
GACAGTAGAC	CCTGACACAC	ACACACAAAA	AAAAACCTTC	ATAAAAAATT	ATTAGTTGAC	1118
TTTTCTTAGG	TGACTTTCCG	TTTAAGCAAT	AAATTTAAAA	GTAAAATCTC	TAATTTTAGA	1178
AAATTTATTT	TTAGTTACAT	ATTGAAATTT	TTAAACCCTA	GGTTTAAGTT	TTATGTCTAA	1238
ATTACCTGAG	AACACACTAA	GTCTGATAAG	CTTCATTTTA	TGGGCCTTTT	GGATGATTAT	1298
ATAATATTCT	GATGAAAGCC	AAGACAGACC	CTTAAACCAT	AAAAATAGGA	GTTTCGAGAA	1358
GAGGAGTAGC	AAAAGTAAAA	GCTAGAATGA	GATTGAATTC	TGAGTCGAAA	TACAAAATTT	1418
TACATATTCT	GTTTCTCTCT	TTTTCCCCCT	CTTAG	CT	GAA GAT GAT G	1470

Ala Glu Asp Asp Glu

-10

GTAGAAATGA	ATTTATTTTT	CTTTGCAAAC	TAAGTATCTG	CTTGAGACAC	ATCTATCTCA	1530
CCATTGTCAG	CTGAGGAAAA	AAAAAAATGG	TTCTCATGCT	ACCAATCTGC	CTTCAAAGAA	1590
ATGTGGACTC	AGTAGCACAG	CTTTGGAATG	AAGATGATCA	TAAGAGATAC	AAAGAAGAAC	1650
CTCTAGCAAA	AGATGCTTCT	CTATGCCTTA	AAAAATTCTC	CAGCTCTTAG	AATCTACAAA	1710
ATAGACTTTG	CCTGTTTTAT	TGGTCCTAAG	ATTAGCATGA	AGCCATGGAT	TCTGTTGTAG	1770
GGGGAGCGTT	GCATAGGAAA	AAGGGATTGA	AGCATTAGAA	TTGTCCAAAA	TCAGTAACAC	1830
CTCCTCTCAG	AAATGCTTTG	GGAAGAAGCC	TGGAAGGTTT	CGGGTTGGTG	GTGGGGTGGG	1890
GCAGAAAATT	CTGGAAGTAG	AGGAGATAGG	AATGGGTGGG	GCAAGAAGAC	CACATTCAGA	1950
GGCCAAAAGC	TGAAAGAAAC	CATGGCATT	ATGATGAATT	CAGGGTAATT	CAGAATGGAA	2010
GTAGAGTAGG	AGTAGGAGAC	TGGTGAGAGG	AGCTAGAGTG	ATAAACAGGG	TGTAGAGCAA	2070
GACGTTCTCT	CACCCCAAGA	TGTGAAATTT	GGACTTTATC	TTGGAGATAA	TAGGGTTAAT	2130
TAAGCACAAT	ATGTATTAGC	TAGGGTAAAG	ATTAGTTTGT	TGTAACAAAG	ACATCCAAAG	2190
ATACAGTAGC	TGAATAAGAT	AGAGAAATTT	TCTCTCAAAG	AAAGTCTAAG	TAGGCAGCTC	2250
AGAAGTAGTA	TGGCTGGAAG	CAACCTGATG	ATATTGGGAC	CCCCAACCTT	CTTCAGTCTT	2310
GTATCCCATCA	TCCCCTAGTT	GTTGATCTCA	CTCACATAGT	TGAAAATCAT	CATACTTCCT	2370
GGGTTTCATAT	CCCAGTTATC	AAGAAAGGGT	CAAGAGAAGT	CAGGCTCAT	CCTTTCAAAG	2430
ACTCTAATTG	GAAGTTAAAC	ACATCAATCC	CCCTCATATT	CCATTGACTA	GAATTTAATC	2490
ACATGGCCAC	ACCAAGTGCA	AGGAAATCTG	GAAAATATAA	TCTTTATTCC	AGGTAGCCAT	2550
ATGACTCTTT	AAAATTCAGA	AATAATATAT	TTTTAAAATA	TCATTCTGGC	TTTGGTATAA	2610
AGAATTGATG	GTGTGGGGTG	AGGAGGCCAA	AATTAAGGGT	TGAGAGCCTA	TTATTTTAGT	2670
TATTACAAGA	AATGATGGTG	TCATGAATTA	AGGTAGACAT	AGGGGAGTGC	TGATGAGGAG	2730
CTGTGAATGG	ATTTTAGAAA	CACCTGAGAG	AATCAATAGG	ACATGATTTA	GGGTTGGATT	2790
TGGAAAGGAG	AAGAAAGTAG	AAAAGATGAT	GCCTACATTT	TTCACCTAGG	CAATTTGTAC	2850
CATTCAAGTGA	AATAGGGAAC	ACAGGAGGAA	GAGCAGGTTT	TGGTGTATAC	AAAGAGGAGG	2910
ATGGATGACG	CATTTTCGTT	TGGATCTGAG	ATGTCTGTGG	AACGTCCTAG	TGGAGATGTC	2970
CACAACTCT	TCTACATGTG	GTTCTGAGTT	CAGGACACAG	ATTTGGGCTG	GAGATAGAGA	3030
TATTGTAGGC	TTATACATAG	AAATGGCATT	TGAATCTATA	GAGATAAAAA	GACACATCAG	3090
AGGAAATGTG	TAAAGTGAGA	GAGGAAAAGC	CAAGTACTGT	GCTGGGGGGA	ATACCTACAT	3150
TTAAAGGATG	CAGTAGAAAG	AAGCTAATAA	ACAACAGAGA	GCAGACTAAC	CAAAAGGGGA	3210
GAAGAAAAAC	CAAGAGAATT	CCACCGACTC	CCAGGAGAGC	ATTTCAAGAT	TGAGGGGATA	3270
GGTGTGTGTG	TGAATTTTGC	AGCCTTGAGA	ATCAAGGGCC	AGAACACAGC	TTTTAGATTT	3330
AGCAACAAGG	AGTTTGGTGA	TCTCAGTGAA	AGCAGCTTGA	TGGTGAAATG	GAGGCAGAGG	3390
CAGATTGCAA	TGAGTGAAAC	AGTGAATGGG	AAGTGAAGAA	ATGATACAGA	TAATTTCTGC	3450
TAAAAGCTTG	GCTGTTAAAA	GGAGGAGAGA	AACAAGACTA	GCTGCAAAGT	GAGATTGGGT	3510
TGATGGAGCA	GTTTTAAATC	TCAAAATAAA	GAGCTTTGTG	CTTTTTTGAT	TATGAAAATA	3570

ATGTGTTAAT	TGTAACATAAT	TGAGGCAATG	AAAAAAGATA	ATAATATGAA	AGATAAAAAAT	3630
ATAAAAACCA	CCCAGAAATA	ATGATAGCTA	CCATTTTGTG	ACAATATTTT	TACACTCCTT	3690
TCTATGTATA	TATACAGACA	CAGAAATGCT	TATATTTT	TTAAAAGGGA	TTGTACTATA	3750
CCTAAGCTGC	TTTTTCTAGT	TAGTGATATA	TATGGACATC	TCTCCATGGC	AACGAGTAAT	3810
TGCAGTTATA	TTAAGTTCAT	GATATTTT	AATAAGGGCA	TATCTTTGCC	CTTTTTATTT	3870
AATCAATTCT	TAATTGGTGA	ATGTTTGT	CCAGTTTGT	GTTGTTATTA	ACAATGTTCC	3930
CATAAGCATT	CCTGTACACC	AATGTTTACA	CATTTGTCTG	ATTTTTTCTT	CAGGATAAAA	3990
CCCAGGAGGT	AGAATTGCTG	GGTTGATAGA	AGAGAAAGGA	TGATTGCCAA	ATTAAAGCTT	4050
CAGTAGAGGG	TACATGCCGA	GCACAAATGG	GATCAGCCCT	AGATAACCAGA	AATGGCACTT	4110
TCTCATTTCC	ECTTGGGACA	AAAGGGAGAG	AGGCAATAAC	TGTGCTGCCA	GAGTTAAATT	4170
TGTACGTGGA	GTAGCAGGAA	ATCATTTGCT	GAAAATGAAA	ACAGAGATGA	TGTTGTAGAG	4230
GTCCTGAAGA	GAGCAAAGAA	AATTTGAAAT	TGCGGCTATC	AGCTATGGAA	GAGAGTGCTG	4290
AACTGGAAAA	CAAAAGAAGT	ATTGACAATT	GGTATGCTTG	TAATGGCACC	GATTTGAACG	4350
CTTGTGCCAT	TGTTTACCAG	CAGCACTCAG	CAGCCAAGTT	TGGAGTTTTG	TAGCAGAAAG	4410
ACAAATAAGT	TAGGGATTTA	ATATCCTGGC	CAAATGGTAG	ACAAAATGAA	CTCTGAGATC	4470
CAGCTGCACA	GGGAAGGAAG	GGAAGACGGG	AAGAGGTTAG	ATAGGAAATA	CAAGAGTCAG	4530
GAGACTGGAA	GATGTTGTGA	TATTTAAGAA	CACATAGAGT	TGGAGTAAAA	GTGTAAGAAA	4590
ACTAGAAGGG	TAAGAGACCG	GTCAGAAAGT	AGGCTATTTG	AAGTTAACAC	TTCAGAGGCA	4650
GAGTAGTTCT	GAATGGTAAC	AAGAAATTGA	GTGTGCCCTT	GAGAGTAGGT	TAAAAAACAA	4710
TAGGCAACTT	TATTGTAGCT	ACTTCTGGAA	CAGAAGATTG	TCATTAATAG	TTTTAGAAAA	4770
CTAAAATATA	TAGCATACTT	ATTTGTCAAT	TAACAAAGAA	ACTATGTATT	TTTAAATGAG	4830
ATTTAATGTT	TATTGTAG	AA AAC CTG	GAA TCA GAT	TAC TTT GGC	AAG CTT	4880
		Glu Asn Leu	Glu Ser Asp	Tyr Phe Gly	Lys Leu	
		-5		1	5	
GAA TCT AAA TTA TCA GTC ATA AGA AAT TTG AAT	GAC CAA GTT CTC TTC					4928
Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn	Asp Gln Val Leu Phe					
		10	15	20		
ATT GAC CAA GGA AAT CGG CCT CTA TTT GAA GAT	ATG ACT GAT TCT GAC					4976
Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp	Met Thr Asp Ser Asp					
		25	30	35		
TGT AGA G GTATTTTTT TTAATTCGCA AACATAGAAA	TGACTAGCTA CTTCTTCCCA					5032
Cys Arg Asp						
		40				
TTCTGTTTTA CTGCTTACAT TGTTCCTGTC TAGTCCCAAT	CCTCAGATGA AAAGTCAACAG					5092
GAGTGACAAAT AATTTCACCT ACAGGAAACT TTATAAGGCA	TCCACGTTTTT TTAGTTGGGG					5152
TAAAAAATTG GATACAATAA GACATTGCTA GGGGTCATGC	CTCTCTGAGC CTGCCTTTGA					5212
ATCACCAATC CCTTTATTGT GATTGCATTA ACTGTTTAAA	ACCTCTATAG TTGGATGCTT					5272
AATCCCTGCT TGTTACAGCT GAAAATGCTG ATAGTTTACC	AGGTGTGGTG GCATCTATCT					5332
GTAATCCTAG CTACTTGGGA GGCTCAAGCA GGAGGATTGC	TTGAGGCCAG GACTTTGAGG					5392
CTGTAGTACA CTGTGATCGT ACCTGTGAAT AGCCACTGCA	CTCCAGCCTG GGTGATATAC					5452
AGACCTTGTC TCTAAAATTA AAAAAAAAAA AAAAAAAAAAC	CTTAGGAAAG GAAATTGATC					5512
AAGTCTACTG TGCCTTCCAA AACATGAATT CCAAATATCA	AAGTTAGGCT GAGTTGAAGC					5572
AGTGAATGTG CATTCCTTAA AAATACTGAA TACTTACCTT	AACATATATT TTAAATATTT					5632
TATTTAGCAT TTAAAAGTTA AAAACAATCT TTTAGAATTC	ATATCTTTAA AATACTCAAA					5692
AAAGTTGCAG CGTGTGTGTT GTAATACACA TTAAACTGTG	GGGTTGTTTG TTTGTTTGAG					5752
ATGCAGTTTC ACTCTGTCAC CCAGGCTGAA GTGCAGTGCA	GTGCAGTGGT GTGATCTCGG					5812
CTCACTACAA CCTCCACCTC CCACGTTCAA GCGATTCTCA	TGCCTCAGTC TCCCGAGTAG					5872
GTGGGATTAC AGGCATGCAC CACTTACACC CGGCTAATTT	TTGTATTTTT AGTAGAGCTG					5932
GGGTTTCACC ATGTTGGCCA GGCTGGTCTC AAACCCCTAA	CCTCAAGTGA TCTGCCTGCC					5992
TCAGCCTCCC AAACAAACAA ACAACCCAC AGTTTAAATAT	GTGTTACAAC ACACATGCTG					6052
CAACTTTTAT GAGTATTTTA ATGATATAGA TTATAAAAGG	TTGTTTTTAA CTTTTAAATG					6112
CTGGGATTAC AGGCATGAGC CACTGTGCCA GGCCTGAACT	GTGTTTTTAA AAATGTCTGA					6172
CCAGCTGTAC ATAGTCTCCT GCAGACTGGC CAAGTCTCAA	AGTGGGAACA GGTGTATTAA					6232
GGACTATCCT TTGGTTAAAT TTCCGCAAAT GTTCTGTGTC	AAGAATTCTT CTAAGTAGAG					6292
TTCTCATTTA TTATATTTAT TTCAG AT AAT GCA CCC	CGG ACC ATA TTT ATT					6343
	Asp Asn Ala Pro Arg Thr Ile Phe Ile					

							40					45					
ATA	AGT	ATG	TAT	AAA	GAT	AGC	CAG	CCT	AGA	GGT	ATG	GCT	GTA	ACT	ATC		6391
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile		
	50					55					60						
TCT	GTG	AAG	TGT	GAG	AAA	ATT	TCA	ACT	CTC	TCC	TGT	GAG	AAC	AAA	ATT		6439
Ser	Val	Lys	Cys	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Cys	Glu	Asn	Lys	Ile		
65					70					75				80			
ATT	TCC	TTT	AAG	GTAAG	ACTGAGCCTT	ACTTTGTTTT	CAATCATGTT	AATATAATCA									6496
Ile	Ser	Phe	Lys														
ATATAATTAG	AAATATAACA	TTATTTCTAA	TGTTAATATA	AGTAATGTAA	TTAGAAAAC												6556
CAATATCCT	CAGACCAACC	TTTTGTCTAG	AACAGAAATA	ACAAGAAGCA	GAGAACCATT												6616
AAAGTGAATA	CTTACTAAAA	ATTATCAAAC	TCTTTACCTA	TTGTGATAAT	GATGGTTTTT												6676
CTGAGCCTGT	CACAGGGGAA	GAGGAGATAC	AACACTTGTT	TTATGACCTG	CATCTCCTGA												6736
ACAATCAGTC	TTTATACAAA	TAATAATGTA	GAATACATAT	GTGAGTTATA	CATTTAAGAA												6796
TAACATGTGA	CTTTCCAGAA	TGAGTTCTGC	TATGAAGAAT	GAAGCTAATT	ATCCTTCTAT												6856
ATTTCTACAC	CTTTGTAAAT	TATGATAATA	TTTTAATCCC	TAGTTGTTTT	GTTGCTGATC												6916
CTTAGCCTAA	GTCTTAGACA	CAAGCTTCAG	CTTCCAGTTG	ATGTATGTTA	TTTTTAATGT												6976
TAATCTAATT	GAATAAAAGT	TATGAGATCA	GCTGTAAAAG	TAATGCTATA	ATTATCTTCA												7036
AGCCAGGTAT	AAAGTATTTC	TGGCCTCTAC	TTTTTCTCTA	TTATTCTCCA	TTATTATTCT												7096
CTATTATTTT	TCTCTATTTC	CTCCATTATT	GTTAGATAAA	CCACAATTAA	CTATAGCTAC												7156
AGACTGAGCC	AGTAAGAGTA	GCCAGGGATG	CTTACAAATT	GGCAATGCTT	CAGAGGAGAA												7216
TTCCATGTCA	TGAAGACTCT	TTTTGAGTGG	AGATTTGCCA	ATAAATATCC	GCTTTCATGC												7276
CCACCCAGTC	CCCCTGAAA	GACAGTTAGG	ATATGACCTT	AGTGAAGGTA	CCAAGGGGCA												7336
ACTTGGTAGG	GAGAAAAAAG	CCACTCTAAA	ATATAATCCA	AGTAAGAACA	GTGCATATGC												7396
AACAGATACA	GCCCCCAGAC	AAATCCCTCA	GCTATCTCCC	TCCAACCAGA	GTGCCACCCC												7456
TTCAGGTGAC	AATTTGGAGT	CCCCATTCTA	GACCTGACAG	GCAGCTTAGT	TATCAAAATA												7516
GCATAAGAGG	CCTGGGATGG	AAGGGTAGGG	TGGAAAGGGT	TAAGCATGCT	GTTACTGAAC												7576
AACATAATTA	GAAGGGAAGG	AGATGGCCAA	GCTCAAGCTA	TGTGGGATAG	AGGAAAAC												7636
AGCTGCAGAG	GCAGATTTCAG	AAACTGGGAT	AAGTCCGAAC	CTACAGGTGG	ATTCTTGTTG												7696
AGGGAGACTG	GTGAAAATGT	TAAGAAGATG	GAAATAATGC	TTGGCACTTA	GTAGGAAC												7756
GGCAAATCCA	TATTTGGGGG	AGCCTGAAGT	TTATTCAATT	TTGATGGCCC	TTTTAAATAA												7816
AAAGAATGTG	GCTGGGCGTG	GTGGCTCACA	CCTGTAATCC	GAGCACTTTG	GGAGGCCGAG												7876
GGGGGCGGAT	CACCTGAAGT	CAGGAGTTCA	AGACCAGCCT	GACCAACATG	GAGAAACCCC												7936
ATCTCTACTA	AAAATACAAA	ATTAGCTGGG	CGTGGTGCCA	TATGCCTGTA	ATCCCAGCTA												7996
CTCGGGAGGC	TGAGGCAGGA	GAATCTTTTG	AACCCGGGAG	GCAGAGGTTG	CGATGAGCCT												8056
AGATCGTGCC	ATTGCACTCC	AGCCTGGGCA	ACAAGAGCAA	AACTCGGTCT	CAAAAAAAAAA												8116
AAAAAAAAAAG	TGAAATTAAC	CAAAGGCATT	AGCTTAATAA	TTTAATACTG	TTTTTAAGTA												8176
GGGCGGGGGG	TGGCTGGAAG	AGATCTGTGT	AAATGAGGGA	ATCTGACATT	TAAGCTTCAT												8236
CAGCATCATA	GCAAATCTGC	TTCTGGAAGG	AACTCAATAA	ATATTAGTTG	GAGGGGGGGA												8296
GAGAGTGAGG	GGTGGACTAG	GACCAGTTTT	AGCCCTTGTC	TTTAATCCCT	TTTCCTGCCA												8356
CTAATAAGGA	TCTTAGCAGT	GGTTATAAAA	GTGGCCTAGG	TTCTAGATAA	TAAGATACAA												8416
CAGGCCAGGC	ACAGTGGCTC	ATGCCTATAA	TCCCAGCACT	TTGGGAGGGC	AAGGCGAGTG												8476
TCTCACTTGA	GATCAGGAGT	TCAAGACCAG	CCTGGCCAGC	ATGGCGATAC	TCTGTCTCTA												8536
CTAAAAAAAAA	TACAAAAATT	AGCCAGGCAT	GGTGGCATGC	ACCTGTAATC	CCAGCTACTC												8596
GTGAGCCTGA	GGCAGAAGAA	TCGCTTGAAA	CCAGGAGGTG	TAGGCTGCAG	TGAGCTGAGA												8656
TCGCACCACT	GCACTCCAGC	CTGGGCGACA	GAATGAGACT	TTGTCTCAAA	AAAAGAAAAA												8716
GATACAACAG	GCTACCCTTA	TGTGCTCACC	TTTCACTGTT	GATTACTAGC	TATAAAGTCC												8776
TATAAAGTTC	TTTGGTCAAG	AACCTTGACA	ACACTAAGAG	GGATTGCTT	TGAGAGGTTA												8836
CTGTCAGAGT	CTGTTTCATA	TATATACATA	TACATGTATA	TATGTATCTA	TATCCAGGCT												8896
TGGCCAGGGT	TCCCTCAGAC	TTTCCAGTGC	ACTTGGGAGA	TGTTAGGTCA	ATATCAACTT												8956
TCCCTGGATT	CAGATTCAAC	CCCTTCTGAT	GTAATAAAAA	AAAAAAAAAA	GAAAGAAATC												9016
CCTTTCCCTT	TGGAGCACTC	AAGTTTCACC	AGGTGGGGCT	TTCCAAGTTG	GGGGTTCTCC												9076
AAGGTCATTG	GGATTGCTTT	CACATCCATT	TGCTATGTAC	CTTCCCTATG	ATGGCTGGGA												9136
GTGGTCAACA	TCAAACTAG	GAAAGCTACT	GCCCAAGGAT	GTCTTACCT	CTATTCTGAA												9196
ATGTGCAATA	AGTGTGATTA	AAGAGATTGC	CTGTTCTACC	TATCCACACT	CTCGCTTTCA												9256
ACTGTAACCT	TCTTTTTTTC	TTTTTTTCTT	TTTTTCTTTT	TTTTTGAAAC	GGAGTCTCGC												9316

TCTGTCGCCC	AGGCTAGAGT	GCAGTGGCAC	GATCTCAGCT	CACTGCAAGC	TCTGCCTCCC	9376
GGGTTTCACGC	CATTCTCCTG	CCTCACCCCTC	CCAAGCAGCT	GGGACTACAG	GCGCCTGCCA	9436
CCATGCCCCAG	CTAATTTTTT	GTATTTTTTAG	TAGAGACGGG	GTTTCACCGT	GTTAGCCAGG	9496
ATGGTCTCGA	TCTCCTGAAC	TTGTGATCCG	CCCGCCTCAG	CCTCCCAAAG	TGCTGGGATT	9556
ACAGGCGTGA	GCCATCGCAC	CCGGCTCAAC	TGTAACCTTTC	TATACTGGTT	CATCTTCCCC	9616
TGTAATGTTA	CTAGAGCTTT	TGAAGTTTTG	GCTATGGATT	ATTTCTCATT	TATACATTAG	9676
ATTTTCAGATT	AGTTCCAAAT	TGATGCCCCAC	AGCTTAGGGT	CTCTTCCTAA	ATTGTATATT	9736
GTAGACAGCT	GCAGAAAGTGG	GTGCCAATAG	GGGAACTAGT	TTATACTTTC	ATCAACTTAG	9796
GACCCACACT	TGTTGATAAA	GAACAAAGGT	CAAGAGTTAT	GACTACTGAT	TCCACAACCTG	9856
ATTGAGAAGT	TGGAGATAAC	CCCGTGACCT	CTGCCATCCA	GAGTCTTTCA	GGCATCTTTG	9916
AAGGATGAAG	AAATGCTATT	TTAATTTTTG	AGGTTTCTCT	ATCAGTGCTT	AGGATCATGG	9976
GAATCTGTGC	TGCCATGAGG	CCAAAATTAA	GTCCAAAACA	TCTACTGGTT	CCAGGATTAA	10036
CATGGAAGAA	CCTTAGGTGG	TGCCCCACATG	TTCTGATCCA	TCCTGCAAAA	TAGACATGCT	10096
GCACTAACAG	GAAAAGTGCA	GGCAGCACTA	CCAGTTGGAT	AACCTGCAAG	ATTATAGTTT	10156
CAAGTAATCT	AACCATTTCT	CACAAGGCC	TATTCTGTGA	CTGAAACATA	CAAGAATCTG	10216
CATTTGGCCT	TCTAAGGCAG	GGCCCAGCCA	AGGAGACCAT	ATTCAGGACA	GAAATTCAAG	10276
ACTACTATGG	AACTGGAGTG	CTTGGCAGGG	AAGACAGAGT	CAAGGACTGC	CAACTGAGCC	10336
AATACAGCAG	GCTTACACAG	GAACCCAGGG	CCTAGCCCTA	CAACAATTAT	TGGGTCTATT	10396
CACTGTAAGT	TTTAATTTCA	GGCTCCACTG	AAAGAGTAAG	CTAAGATTCC	TGGCACTTTC	10456
TGTCTCTCTC	ACAGTTGGCT	CAGAAATGAG	AACTGGTCAG	GCCAGGCATG	GTGGCTTACA	10516
CCTGGAATCC	CAGCACTTTG	GGAGGCCGAA	GTGGGAGGGT	CACTTGAGGC	CAGGAGTTCA	10576
GGACCAGCTT	AGGCAACAAA	GTGAGATACC	CCCTGACCCC	TTCTCTACAA	AAATAAATTT	10636
TAAAAATTAG	CCAAATGTGG	TGGTGTATAC	TTACAGTCCC	AGCTACTCAG	GAGGCTGAGG	10696
CAGGGGGATT	GCTTGAGCCC	AGGAATTCAA	GGCTGCAGTG	AGCTATGATT	TCACCACTGC	10756
ACTTCTGGCT	GGGCAACAGA	GCGAGACCCT	GTCTCAAAGC	AAAAAGAAAA	AGAAACTAGA	10816
ACTAGCCTAA	GTTTGTGGGA	GGAGGTCATC	ATCGTCTTTA	CCCGTGAATG	GTTATTATAG	10876
AGGACAGAAA	TTGACATTAG	CCCAAAAAGC	TTGTGGTCTT	TGCTGGAACT	CTACTTAATC	10936
TTGAGCAAAT	GTGGACACCA	CTCAATGGGA	GAGGAGAGAA	GTAAGCTGTT	TGATGTATAG	10996
GGGAAAACCTA	GAGGCCTGGA	ACTGAATATG	CATCCCATGA	CAGGGAGAAT	AGGAGATTCTG	11056
GAGTTAAGAA	GGAGAGGAGG	TCAGTACTGC	TGTTTCAGAGA	TTTTTTTTTAT	GTAACCTCTTG	11116
AGAAGCAAAA	CTACTTTTGT	TCTGTTTGGT	AATATACTTC	AAAACAAACT	TCATATATTC	11176
AAATTGTTCA	TGTCCTGAAA	TAATTAGGTA	ATGTTTTTTT	CTCTATAG	GAA ATG AAT	11233

Glu Met Asn
85

CCT	CCT	GAT	AAC	ATC	AAG	GAT	ACA	AAA	AGT	GAC	ATC	ATA	TTC	TTT	CAG	11281
Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys	Ser	Asp	Ile	Ile	Phe	Phe	Glu	
		90					95					100				
AGA	AGT	GTC	CCA	GGA	CAT	GAT	AAT	AAG	ATG	CAA	TTT	GAA	TCT	TCA	TCA	11329
Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys	Met	Gln	Phe	Glu	Ser	Ser	Ser	
		105				110					115					
TAC	GAA	GGA	TAC	TTT	CTA	GCT	TGT	GAA	AAA	GAG	AGA	GAC	CTT	TTT	AAA	11377
Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Cys	Glu	Lys	Glu	Arg	Asp	Leu	Phe	Lys	
		120			125					130					135	
CTC	ATT	TTG	AAA	AAA	GAG	GAT	GAA	TTG	GGG	GAT	AGA	TCT	ATA	ATG	TTC	11425
Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu	Gly	Asp	Arg	Ser	Ile	Met	Phe	
			140					145						150		
ACT	GTT	CAA	AAC	GAA	GAC	TAGCTATTAA	AATTTTCATGC	C								11464
Thr	Val	Gln	Asn	Glu	Asp											
			155													

(18) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 471 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: cDNA to mRNA

(vi)ORIGINAL SOURCE:

(A)ORGANISM: mouse

(G)CELL TYPE: liver

(ix)FEATURE:

(A)NAME/KEY: mat peptide

(B)LOCATION: 1..471

(C)IDENTIFICATION METHOD: S

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 18:

AAC	TTT	GGC	CGA	CTT	CAC	TGT	ACA	ACC	GCA	GTA	ATA	CGG	AAT	ATA	AAT	48
Asn	Phe	Gly	Arg	Leu	His	Cys	Thr	Thr	Ala	Val	Ile	Arg	Asn	Ile	Asn	
1				5					10					15		
GAC	CAA	GTT	CTC	TTC	GTT	GAC	AAA	AGA	CAG	CCT	GTG	TTC	GAG	GAT	ATG	96
Asp	Gln	Val	Leu	Phe	Val	Asp	Lys	Arg	Gln	Pro	Val	Phe	Glu	Asp	Met	
			20					25					30			
ACT	GAT	ATT	GAT	CAA	AGT	GCC	AGT	GAA	CCC	CAG	ACC	AGA	CTG	ATA	ATA	144
Thr	Asp	Ile	Asp	Gln	Ser	Ala	Ser	Glu	Pro	Gln	Thr	Arg	Leu	Ile	Ile	
		35					40					45				
TAC	ATG	TAC	AAA	GAC	AGT	GAA	GTA	AGA	GGA	CTG	GCT	GTG	ACC	CTC	TCT	192
Tyr	Met	Tyr	Lys	Asp	Ser	Glu	Val	Arg	Gly	Leu	Ala	Val	Thr	Leu	Ser	
	50					55				60						
GTG	AAG	GAT	AGT	AAA	ATG	TCT	ACC	CTC	TCC	TGT	AAG	AAC	AAG	ATC	ATT	240
Val	Lys	Asp	Ser	Lys	Met	Ser	Thr	Leu	Ser	Cys	Lys	Asn	Lys	Ile	Ile	
65					70				75					80		
TCC	TTT	GAG	GAA	ATG	GAT	CCA	CCT	GAA	AAT	ATT	GAT	GAT	ATA	CAA	AGT	288
Ser	Phe	Glu	Glu	Met	Asp	Pro	Pro	Glu	Asn	Ile	Asp	Asp	Ile	Gln	Ser	
				85					90					95		
GAT	CTC	ATA	TTC	TTT	CAG	AAA	CGT	GTT	CCA	GGA	CAC	AAC	AAG	ATG	GAG	336
Asp	Leu	Ile	Phe	Phe	Gln	Lys	Arg	Val	Pro	Gly	His	Asn	Lys	Met	Glu	
			100					105					110			
TTT	GAA	TCT	TCA	CTG	TAT	GAA	GGA	CAC	TTT	CTT	GCT	TGC	CAA	AAG	GAA	384
Phe	Glu	Ser	Ser	Leu	Tyr	Glu	Gly	His	Phe	Leu	Ala	Cys	Gln	Lys	Glu	
		115					120					125				
GAT	GAT	GCT	TTC	AAA	CTC	ATT	CTG	AAA	AAA	AAG	GAT	GAA	AAT	GGG	GAT	432
Asp	Asp	Ala	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Lys	Asp	Glu	Asn	Gly	Asp	
	130					135					140					
AAA	TCT	GTA	ATG	TTC	ACT	CTC	ACT	AAC	TTA	CAT	CAA	AGT				471
Lys	Ser	Val	Met	Phe	Thr	Leu	Thr	Asn	Leu	His	Gln	Ser				
145					150					155						

(19) INFORMATION FOR SEQ ID NO: 19:

(i)SEQUENCE CHARACTERISTICS:

(A)LENGTH: 9 amino acids

(B)TYPE: amino acid

(D)TOPOLOGY: linear

(ii)MOLECULE TYPE: peptide

(v)FRAGMENT TYPE: N-terminal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Asn Phe Gly Arg Leu His Cys Thr Thr
1 5

(20) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 157 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn
1 5 10 15
Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp
20 25 30
Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile
35 40 45
Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile
50 55 60
Ser Val Lys Ser Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile
65 70 75 80
Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys
85 90 95
Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys
100 105 110
Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu
115 120 125
Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu
130 135 140
Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp
145 150 155

(21) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 157 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn
1 5 10 15
Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp
20 25 30
Met Thr Asp Ser Asp Ser Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile
35 40 45
Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile
50 55 60

Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Cys	Glu	Asn	Lys	Ile
65					70					75					80
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys
			85						90					95	
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
			100					105					110		
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Cys	Glu
		115					120					125			
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu
	130					135					140				
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp			
145					150					155					

(22) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn
1			5					10						15	
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp
		20					25					30			
Met	Thr	Asp	Ser	Asp	Cys	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile
		35				40					45				
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile
	50				55					60					
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Cys	Glu	Asn	Lys	Ile
65				70					75					80	
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys
			85					90					95		
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
			100					105				110			
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Ser	Glu
		115					120					125			
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu
	130					135					140				
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp			
145					150					155					

(23) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn
1				5					10					15	
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp
			20					25					30		
Met	Thr	Asp	Ser	Asp	Ser	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile
		35					40					45			
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile
	50					55					60				
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Cys	Glu	Asn	Lys	Ile
65					70					75				80	
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys
				85					90					95	
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
			100					105					110		
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Ser	Glu
		115					120					125			
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu
	130					135					140				
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp			
145					150					155					

(24) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn
1				5					10					15	
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp
			20					25					30		
Met	Thr	Asp	Ser	Asp	Ser	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile
		35					40					45			
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile
	50					55					60				
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Ser	Glu	Asn	Lys	Ile
65					70					75				80	
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys
				85					90					95	
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
			100					105					110		
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Ser	Glu
		115					120					125			
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu
	130					135					140				
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp			
145					150					155					

(25) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 157 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn
1				5					10					15	
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp
			20					25					30		
Met	Thr	Asp	Ser	Asp	Ser	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile
		35					40					45			
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile
	50					55					60				
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Ala	Glu	Asn	Lys	Ile
65					70					75				80	
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys
				85					90					95	
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
			100					105					110		
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Cys	Glu
		115					120					125			
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu
	130					135					140				
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp			
145					150					155					

(26) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 157 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Tyr	Phe	Gly	Lys	Leu	Glu	Ser	Lys	Leu	Ser	Val	Ile	Arg	Asn	Leu	Asn
1				5					10					15	
Asp	Gln	Val	Leu	Phe	Ile	Asp	Gln	Gly	Asn	Arg	Pro	Leu	Phe	Glu	Asp
			20					25					30		
Met	Thr	Asp	Ser	Asp	Ser	Arg	Asp	Asn	Ala	Pro	Arg	Thr	Ile	Phe	Ile
		35					40					45			
Ile	Ser	Met	Tyr	Lys	Asp	Ser	Gln	Pro	Arg	Gly	Met	Ala	Val	Thr	Ile
	50					55					60				
Ser	Val	Lys	Ser	Glu	Lys	Ile	Ser	Thr	Leu	Ser	Ala	Glu	Asn	Lys	Ile
65					70					75				80	
Ile	Ser	Phe	Lys	Glu	Met	Asn	Pro	Pro	Asp	Asn	Ile	Lys	Asp	Thr	Lys
				85					90					95	
Ser	Asp	Ile	Ile	Phe	Phe	Gln	Arg	Ser	Val	Pro	Gly	His	Asp	Asn	Lys
			100					105					110		
Met	Gln	Phe	Glu	Ser	Ser	Ser	Tyr	Glu	Gly	Tyr	Phe	Leu	Ala	Ser	Glu
		115					120					125			
Lys	Glu	Arg	Asp	Leu	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Glu	Asp	Glu	Leu

	130		135		140							
Gly	Asp	Arg	Ser	Ile	Met	Phe	Thr	Val	Gln	Asn	Glu	Asp
145					150					155		

(27) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 157 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Asn	Phe	Gly	Arg	Leu	His	Ala	Thr	Thr	Ala	Val	Ile	Arg	Asn	Ile	Asn
1				5					10					15	
Asp	Gln	Val	Leu	Phe	Val	Asp	Lys	Arg	Gln	Pro	Val	Phe	Glu	Asp	Met
			20					25					30		
Thr	Asp	Ile	Asp	Gln	Ser	Ala	Ser	Glu	Pro	Gln	Thr	Arg	Leu	Ile	Ile
		35					40					45			
Tyr	Met	Tyr	Lys	Asp	Ser	Glu	Val	Arg	Gly	Leu	Ala	Val	Thr	Leu	Ser
	50					55					60				
Val	Lys	Asp	Ser	Lys	Met	Ser	Thr	Leu	Ser	Cys	Lys	Asn	Lys	Ile	Ile
65					70					75					80
Ser	Phe	Glu	Glu	Met	Asp	Pro	Pro	Glu	Asn	Ile	Asp	Asp	Ile	Gln	Ser
				85					90					95	
Asp	Leu	Ile	Phe	Gln	Lys	Arg	Val	Pro	Gly	His	Asn	Lys	Met	Glu	
			100				105						110		
Phe	Glu	Ser	Ser	Leu	Tyr	Glu	Gly	His	Phe	Leu	Ala	Cys	Gln	Lys	Glu
		115					120					125			
Asp	Asp	Ala	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Lys	Asp	Glu	Asn	Gly	Asp
	130					135					140				
Lys	Ser	Val	Met	Phe	Thr	Leu	Thr	Asn	Leu	His	Gln	Ser			
145					150					155					

(28) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 157 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Asn	Phe	Gly	Arg	Leu	His	Cys	Thr	Thr	Ala	Val	Ile	Arg	Asn	Ile	Asn
1				5					10					15	
Asp	Gln	Val	Leu	Phe	Val	Asp	Lys	Arg	Gln	Pro	Val	Phe	Glu	Asp	Met
			20					25					30		
Thr	Asp	Ile	Asp	Gln	Ser	Ala	Ser	Glu	Pro	Gln	Thr	Arg	Leu	Ile	Ile
		35					40					45			
Tyr	Met	Tyr	Lys	Asp	Ser	Glu	Val	Arg	Gly	Leu	Ala	Val	Thr	Leu	Ser
	50					55					60				
Val	Lys	Asp	Ser	Lys	Met	Ser	Thr	Leu	Ser	Cys	Lys	Asn	Lys	Ile	Ile

65					70					75					80
Ser	Phe	Glu	Glu	Met	Asp	Pro	Pro	Glu	Asn	Ile	Asp	Asp	Ile	Gln	Ser
				85					90					95	
Asp	Leu	Ile	Phe	Phe	Gln	Lys	Arg	Val	Pro	Gly	His	Asn	Lys	Met	Glu
			100					105					110		
Phe	Glu	Ser	Ser	Leu	Tyr	Glu	Gly	His	Phe	Leu	Ala	Ser	Gln	Lys	Glu
		115					120					125			
Asp	Asp	Ala	Phe	Lys	Leu	Ile	Leu	Lys	Lys	Lys	Asp	Glu	Asn	Gly	Asp
	130					135					140				
Lys	Ser	Val	Met	Phe	Thr	Leu	Thr	Asn	Leu	His	Gln	Ser			
145					150					155					